



جامعة الموصل

كلية علوم الحاسوب والرياضيات

مقارنة بين بعض طرائق تصنيف الانماط باستخدام الدالة التمييزية مع التطبيق

نهى أحمد مجيد المختار

رسالة الدبلوم العالي

الإحصاء

بإشراف

الأستاذ المساعد

د. مثنى صبحي سليمان

١٤٣٩ هـ

٢٠١٨ م

المستخلص

تتناول هذه الرسالة المقارنة بين طرائق مختلفة لتصنيف الانماط والتي تشمل التعلم باشراف والتعلم من دون اشراف، والتي تعتمد بالاساس على المسافات الاقليدية في عملية التصنيف. وقد تم عرض كل من طريقتي التحليل العنقودي وطريقة الانماط المثالية، وتم اجراء المقارنة بين الطريقتين باستخدام الدالة التمييزية. كما تم اجراء دراسة تطبيقية في احد المجالات الطبية المهمة وهو تصنيف مرضى الثلاثيميا نوع بيتا الى صنفين، ثم بيان مدى كفاءة الاصناف التي تم تكوينها من هذه الطرائق، واعطاء الاهمية النسبية للسمات الاساسية والتي تمثل المتغيرات الداخلة في عملية التصنيف. وقد تم التوصل من خلال قيم اخطاء التصنيف والخطأ الحقيقي للدالة التمييزية لكل من طريقتي التحليل العنقودي والانماط المثالية، وهو عامل مهم لإثبات كفاءة الدالة التمييزية، ان الدالة التمييزية لمجاميع للأنماط المثالية هي اكثر كفاءة وافضل في التصنيف من الدالة التمييزية لعناقيد التحليل العنقودي، أي ان طريقة الانماط المثالية افضل من طريقة التحليل العنقودي في تصنيف البيانات قيد الدراسة.

**UNIVERSITY OF MOSUL
COLLEGE OF COMPUTER SCIENCES
AND MATHEMATICS**



Comparison of Some Methods of Patterns Classification by Using Discriminant Function with an Application

Nuha Ahmed Majeed AL-Mukhtar

Higher Diploma

Statistics

Supervised by

Dr. Muthanna Subhi Sulaiman

Assistant Prof.

Abstract

This thesis deals with the comparison of different methods of classification of patterns, which include supervised learning and unsupervised learning, which are based primarily on the Euclidean distances in the classification process. Both Cluster analysis and Ideal patterns methods were presented, and the two methods were compared using the Discriminant function. An applied study was carried out in one of the important medical fields: classification of patients with Beta-Thalassemia to two types. It also showed the efficiency of the classes that were formed from these methods and gave the relative importance of the basic features which represent the variables involved in the classification process. The values of the classification errors and the real error of the Discriminant function of both Cluster analysis and Ideal patterns methods, which is an important factor for demonstrating the efficiency of the discriminant function, were obtained. The Discriminant function of groups of Ideal patterns is more efficient and better in classification than the discriminant function of groups of Cluster analysis. Thus Ideal patterns are better than Cluster analysis in the classification of data under study.